AUTOMATIC PATENT CLAIM READER AND COMPUTER-AIDED CLAIM READING METHOD

FIELD OF INVENTION

[0001] The present invention relates to analyzing patents for scope of claims.

BACKGROUND OF THE INVENTION

[0002] Often, parties holding one or more patents or patent applications need to understand the competitive strengths and/or weaknesses of those patents or patent applications for a given context. The context may be strategic such as during business negotiations like licensing or tactical such as during product design.

[0003] It is therefore often necessary for a human being to read one or more patents or patent applications, analyze the patents or patent applications read, and then somehow relate the analysis to other patents, patent applications, products, and/or services.

[0004] Some methodologies for such analysis have been suggested in the prior art, including looking at how many times a patent has been cited by other patents and/or patent applications. This may be useful in certain circumstances but it does not provide a methodology for assessing the strength of the patent, i.e. the scope of its claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] **Fig. 1** is a schematic overview of an exemplary system;

[0006] **Figs. 2a** and **2b** are exemplars of patent claim structures;

[0007] **Fig. 3** is an illustration of limitations in a claim which may be necessary and therefore surplus; and

[0008] **Fig. 4** is a flowchart of an exemplary embodiment of a method of the present invention.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

[0009] In general, throughout this description, if an item is described as implemented in software, it can equally well be implemented as hardware.

[0010] As used herein, "data" is either singular or plural, as the context requires.

[0011] Referring now to **Fig. 1**, system 1 for patent analysis comprises data store 10, computer 20 operatively connected to data store 10, and parsing program 30 (not shown in the figures) executable in computer 20.

[0012] As used herein, data store 10 may be a persistent read/write data store such as a magnetic storage device, an electronic storage device, a hard drive or a rewritable optical medium, a persistent write-once-read-many data store such as a writable optical medium, a non-persistent data store such as random access memory, or the like, or a combination thereof.

[0013] As further used herein, computer 20 may be any suitable computer capable of being operatively connected to data store 10 and of executing parsing program 30, e.g. a personal computer.

[0014] Parsing program 30 is capable of parsing a patent claim into a set of discrete elements, categorizing each element in the set of discrete elements according to a predetermined rule, and storing a set of categorized elements in the data store. The claim may be made available to parsing program 30 in an electronic or optical or other equivalent format.

[0015] Referring now to **Fig. 2a**, patent claim 40 typically comprises preamble 41 and one or more elements 42.

[0016] Preambles are typically not used to limit elements 42. In certain drafting formats, a stylized format is used in which a prior art environment is first described, followed by the improvement to that prior art. This is sometimes referred to as a Jepson claim. Thus, preamble 41 to the Jepson claim clarifies necessary elements 42 which are not related to the strength of

patent claim 40. Other times, preamble 41 merely sets a field for patent claim 40. However, at other times, preamble 41 may define a term, e.g. a member of a keyword set. Therefore, preamble 41 may need to be parsed to determine the scope of patent claim 40.

[0017] Element 42 further describes a limitation which, as will be familiar to those of ordinary skill in the patent drafting arts, further comprises structural and/or functional terminology. Typically, each element 42 is separated from other elements 42 by punctuation such as a semicolon. Further, each element 42 typically begins on a separate line from other elements. In some cases, each element is further numbered or otherwise identified as a separate element. Referring now to **Fig. 2b**, not all claims are written in the clean form illustrated in **Fig. 2a**. Often, patent drafts are faced with the question: 2b or not 2b?

[0018] Parsing program 30 may further be used to identify one or more keyword sets in a parsed claim. As used herein, a "keyword set" may comprise a noun, an adjective and a noun, a verb, an adverb and a verb, or the like. Keyword sets may be used for further analysis of each claim 40 and its preamble 41 and/or elements 42.

[0019] Referring now to **Fig. 3a** and **Fig. 3b**, additionally, not all elements 42 are meaningful to an analysis of claim 40. For example, claim 40a (**Fig. 3a**) claims a semiconductor device, element 42a of which is a substrate. Claim 40b (**Fig. 3b**) claims an equivalent semiconductor device which lacks a substrate element 42. However, the presence or absence of element 42a for the substrate does not impact the scope of coverage of claim 40b as substrates are necessary for semiconductor devices.

[0020] Accordingly, database 12 (**Fig. 1**) may be present to contain data and/or rules which allow parsing program 30 to identify, within a context of claim 40, those elements 42 which are necessary and therefore which do not affect the scope of claim 40. In a preferred

embodiment, structural elements 42 may be analyzed using just the noun portion of that structural element 42 when identifying if that structural element 42 is necessary.

[0021] Elements 42 may additionally be logically paired with other elements 42. For example, "processing a photoresist layer" as an element 42 in a first claim 40 may be logically paired with elements 42 "applying a photoresist layer" and "removing the photoresist layer" for a semiconductor claim 40. Additionally, if a first claim 40 merely recites "removing the photoresist layer" as an element 42, that element 42 may be logically paired to elements 42 "applying a photoresist layer" and "removing the photoresist layer" for a second semiconductor claim 40.

[0022] Database 12 (**Fig. 1**), or, optionally, another database such as database 14 (**Fig. 1**), may contain a database comprising language equivalents useful to correlate a keyword set in a first expression to a keyword set in a second expression. For example, the correlation may relate a keyword set in English to one in Chinese, or may relate terms which are equivalent such as "RAM" with "random access memory" in a computer context. As another example, as an MOS transistor gate oxide is typically thermally grown on a substrate, the verb "formed" may be an equivalent to "thermally grown" for a claim involving an MOS transistor.

[0023] In a preferred embodiment, rating program 32 (not shown in the figures) is also present and executable in computer 20 (**Fig. 1**). Rating program 32 may be capable of assigning a rating weight to each categorized element 42 (**Fig. 2a**). Assignments of weights may be rules-based, e.g. a rule which takes into consideration the number of useful elements 42 and/or the scope of each element 42.

[0024] **In the operation of an exemplary embodiment,** a patent's claims may be analyzed for scope of coverage. Typically, a claim 40 (**Fig. 2a**) is stronger when it has a fewer

number of elements 42 (**Fig. 2a**), or limitations. Further, typically, scope of claim 40 tends to weaken, e.g. become more narrow, with an increase in the number of elements 42 present in that claim 40. An analysis of a patent or patent application, e.g. a patent or application not owned by the analyzing party who wants to compare that patent or application against other patents or applications which may be owned or licensed by the analyzing party, may therefore consider the number of elements 42 present in each of the claims 40 to be analyzed and the scope of each of those elements 42, e.g. according to the meaning of the wording used for those elements 42.

[0025] Referring now to **Fig. 4**, in an exemplary embodiment, claim 40 (**Fig. 2a**) is retrieved step 100, where claim 40 has been rendered into a format parsable by parsing program 20 (**Fig. 2a**) into a computer memory, e.g. data store 10 (**Fig. 2a**). Once retrieved, parsing program 20 may parse claim 40, step 110, into a set of discrete elements 42.

[0026] As used herein, parsing may be by semantic indexing, latent semantic indexing, rules based parsing, free form parsing, or the like, or a combination thereof. For example, parsing may further comprise using synonyms or equivalents from database 12,14 (**Fig. 3**).

[0027] In a preferred embodiment, parsing may further comprise identifying each keyword set in each element 42 where the keyword set comprises a noun, an adjective and a noun, a verb, an adverb and a verb, or the like. For example, nouns are typically present in structural elements 42 and verbs typically present in functional elements 42.

[0028] By way of example, in **Fig. 2a**, the following may be keyword sets that have been parsed: (1) substrate, (2) transistor devices, (3) metal interconnection, and (4) passivation layer.

[0029] Each keyword set may be analyzed to associate a modifier with the keyword set to categorize the keyword set, step 120. As used herein, a "modifier" may be a modifier identifying the keyword set as a necessary keyword set, a modifier identifying the keyword set as a non-

necessary keyword set, or the like. During further analysis, keyword sets with a necessary modifier may be given less weight than other keyword sets for a claim 40. By way of example, in **Fig. 2a**, substrate may be associated with a "necessary" modifier and transistor devices, metal interconnection, and passivation layer associated with a "non-necessary" modifier.

[0030] As described herein above, a predetermined number of keyword sets may be logically paired with at least one other keyword set, e.g. if a first claim 40 merely recites "removing the photoresist layer" as an element 42, that element 42 may be logically paired to elements 42 "applying a photoresist layer" and "removing the photoresist layer" for a second semiconductor claim 40.

[0031] Parsing program 20 may categorize each element 42 in the set of elements according to a predetermined rule. For example, a categorization attribute may be associated with an element 42 such as a necessary attribute, a non-necessary attribute, a useful attribute, a non-useful attribute, a correlation attribute, or the like, or a combination thereof. As used herein, "necessary" means that this element 42 is assumed to be part of each claim of like type, e.g. all semiconductor transistor devices comprise a substrate. A "non-necessary" attribute may mean the opposite, e.g. this is novel or otherwise not always present in such claimed material. "Useful" may mean that this element 42 helps to distinguish its claim 40 over other patents, and "non-useful" may mean the opposite. "Correlation" may mean that this element 42 may be correlated to another element 42, e.g. a synonym from database 12,14.

[0032] As will be understood by those of ordinary skill in the computer arts, the modifiers, attributes, and logical pairings may be accomplished in a variety of equivalent ways, e.g. a field in a database record, an element in an array, use of different tables, and the like, or a combination thereof.

[0033] Additionally, parsed and categorized elements 42 may be assigned a rating weight to each categorized element. In certain embodiments, an element 42 which is modified such as with a numerical adjective is considered weaker than that element 42 without such a numerical adjective, e.g. "a transistor device" is stronger than "a plurality of transistor devices" which is stronger than "three transistor devices" for purposes of analysis of scope.

[0034] Rating may also take into consideration the number of useful and non-useful elements as well as the scope of each element. By way of example and not limitation, a rating may be obtained using a rule such as:

$$Rating = \sum_{i=1}^{N} \left[(Element_i) \times (GSWeighting) \times (NumericalWeighting) \right]$$

where: $Element_i$ is a weight for the i^{th} element 42 of N elements 42, e.g. "1" for a useful element and "0" for a necessary element;

GSWeighting is a factor which reflects the broad nature of the element, e.g. "1" for a genus claim, "2" for a species claim, "3" for a subspecies claim; and

NumericalWeighting is a factor which reflects whether or not a numerical adjective is present for the i^{th} claim, e.g. "1" for no numerical adjective, "10" for the presence of a numerical adjective

In such a weighting, a higher rating would reflect a claim having a narrower scope than a claim with a lower rating. This example is but one way to assign weights.

[0035] Categorizing may further comprise correlating each element 42 with at least one category in a database of categories, e.g. in database 12,14 (**Fig. 1**).

[0036] Analyzed claims 40 may be filtered, e.g. based on the categorized elements, such as by logically marking only those categorized elements 42 which meet a predetermined rule for the filtering. Such rules may include discarding those claims 40 which do not meet a

predetermined rating weight, discarding those claims 40 which do not include predetermined language for an element 42, or the like, or a combination thereof.

[0037] Categorized elements 42, optionally filtered, may then be stored, step 130, e.g. in data store 10 such as in an interrogatable database.

[0038] In an embodiment, a target patent may be analyzed against a portfolio of patents. A portfolio of patents may be initialized in an interrogatable format, e.g. a computer manipulatable format. One or more patents may be selected from the portfolio of patents for analysis. A predetermined set of claims 40 of the selected patent may be parsed into a set of elements 42, e.g. by parsing program 30 and a rating generated for of each parsed claim 40 of the predetermined set of claims 40 of the selected patent using a predetermined weighting rule.

[0039] The rating such as by rating program 32 may generated according to a database of functions. If desired, rated claims 40 may be sorted according to a rate sorting rule such as a sort based on a rating and on a number of elements 42 present in each claim 40 of the patent being analyzed, a sort based on a rating and on a number of elements 42 present in each claim 40 of the selected patent where the elements 42 are further marked as necessary or non-necessary, a sort based on a rating and on a number of elements 42 present in each claim 40 of the patent where the elements 42 are further marked as useful or non-useful, or the like, or a combination thereof.

[0040] Additionally, a predetermined rule may be used to identify a best claim 40 of the predetermined set of claims 40 of the patent, e.g. one with the broadest scope such as with the lowest rating.

[0041] It will be understood that various changes in the details, materials, and arrangements of the parts which have been described and illustrated above in order to explain the

nature of this invention may be made by those skilled in the art without departing from the principle and scope of the invention as recited in the appended claims.